**🚀 Credit Card Transaction Fraud Detection Project**

**📚 Project Overview**

This project focuses on detecting fraudulent credit card transactions using both **Machine Learning** and a **Convolutional Neural Network (CNN)** model. Due to the imbalanced nature of the dataset, special focus was given to **recall**, **F1-score**, and **AUC** metrics instead of just accuracy.

**📂 Dataset Overview**

* ✅ Source: Kaggle - Credit Card Fraud Detection Dataset
* 🧮 Total transactions: **284,807**
* 🚨 Fraudulent transactions: **492** (0.172%)
* 🧩 Features: V1 – V28 (PCA transformed), Time, Amount, and Class (0 or 1)

**✅ Project Pipeline**

* Data Loading & Exploration
* Preprocessing (Scaling, Balancing, Train-Test Split)
* Exploratory Data Analysis (EDA)
* Model Training & Evaluation across **10 different tests**
* Model Comparison and Conclusion

**📊 Model Evaluation Results (All 10 Tests)**

| **🔎 Test** | **Model** | **Accuracy** | **Precision** | **Recall** | **F1-Score** | **ROC-AUC** |
| --- | --- | --- | --- | --- | --- | --- |
| 1️⃣ | Logistic Regression | 99.93% | 85.71% | 57.14% | 68.57% | 0.936 |
| 2️⃣ | Decision Tree | 99.92% | 75.00% | 61.90% | 68.75% | 0.905 |
| 3️⃣ | Random Forest | 99.97% | 91.67% | 71.42% | 80.00% | 0.946 |
| 4️⃣ | XGBoost | 99.97% | 91.66% | 76.19% | 83.33% | **0.971** |
| 5️⃣ | Support Vector Machine (SVM) | 99.93% | 80.00% | 66.66% | 72.72% | 0.920 |
| 6️⃣ | K-Nearest Neighbors | 99.93% | 76.92% | 61.90% | 68.75% | 0.902 |
| 7️⃣ | Naive Bayes | 99.84% | 16.21% | 71.42% | 26.45% | 0.837 |
| 8️⃣ | Gradient Boosting | 99.97% | 91.66% | 76.19% | 83.33% | 0.969 |
| 9️⃣ | LightGBM | 99.97% | 91.66% | 76.19% | 83.33% | 0.968 |
| 🔟 | **Convolutional Neural Network (CNN)** | 99.97% | 90.00% | 76.19% | 82.35% | 0.965 |

✅ **XGBoost, Gradient Boosting, LightGBM, and CNN** models achieved the best balance between recall and precision, making them the most reliable for fraud detection in this project.

**🔎 Key Takeaways**

* ✅ Accuracy alone is misleading for imbalanced datasets; recall and AUC are critical.
* ✅ XGBoost and Gradient Boosting models consistently performed best.
* ✅ The custom-built **CNN model** also showed excellent results, confirming that deep learning can complement traditional ML in fraud detection tasks.
* ⚠ Naive Bayes, while achieving high recall, resulted in excessive false positives.

**🛠️ Technologies Used**

* **Python**
* **Pandas**, **NumPy**
* **Matplotlib**, **Seaborn**
* **Scikit-learn**
* **XGBoost**, **LightGBM**, **Gradient Boosting**
* **TensorFlow** and **Keras** (for CNN model)

**📁 Project Structure**

arduino

CopyEdit

📂 Credit-Card-Transaction-Fraud-Detection-Project

├─ 📒 Credit\_Card\_Fraud\_Detection.ipynb

├─ 📁 dataset (not included for size reasons)

└─ 📄 README.md

**🚀 How to Run Locally**

1. Clone this repository:

bash

CopyEdit

git clone https://github.com/ShaikhBorhanUddin/Credit-Card-Transaction-Fraud-Detection-Project.git

1. Open Credit\_Card\_Fraud\_Detection.ipynb in **Jupyter Notebook** or **Google Colab**
2. Download the dataset from Kaggle and place it in the project directory
3. Run all cells in sequence to reproduce the results

**🤝 Contributing**

Contributions are welcome!  
Please feel free to fork the repo, create a branch, and submit a pull request for improvements or additional models.

**📬 Contact**

* 🔗 [LinkedIn](https://www.linkedin.com/in/shaikh-borhan-uddin/)
* 💻 GitHub: [@ShaikhBorhanUddin](https://github.com/ShaikhBorhanUddin)

⭐ If you like this project, don’t forget to **star** the repository!